

SK



PCT10

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/018,902

DATE: 05/15/2002

TIME: 16:07:11

Input Set : A:\BB1373 USPCT Corrected Seq List.txt

Output Set: N:\CRF3\05152002\J018902.raw

P6

3 <110> APPLICANT: Rebecca E. Cahoon
4 Steven Gutteridge
5 Leslie T. Harvell
6 J. Antoni Rafalski
7 Yong Tao
8 Zude Weng
10 <120> TITLE OF INVENTION: Polynucleotides Encoding Aminolevulinic Acid Biosynthetic
Enzymes
12 <130> FILE REFERENCE: BB-1373
14 <140> CURRENT APPLICATION NUMBER: 10/018,902
C--> 15 <141> CURRENT FILING DATE: 2002-04-11
17 <150> PRIOR APPLICATION NUMBER: 60/146600
18 <151> PRIOR FILING DATE: 1999-07-30
20 <160> NUMBER OF SEQ ID NOS: 30
22 <170> SOFTWARE: Microsoft Office 97
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 312
26 <212> TYPE: DNA
27 <213> ORGANISM: Zea mays
29 <220> FEATURE:
30 <221> NAME/KEY: unsure
31 <222> LOCATION: (30)
32 <223> OTHER INFORMATION: n=a,c,g or t
34 <220> FEATURE:
35 <221> NAME/KEY: unsure
36 <222> LOCATION: (247)
37 <223> OTHER INFORMATION: n=a,c,g or t
39 <220> FEATURE:
40 <221> NAME/KEY: unsure
41 <222> LOCATION: (256)
42 <223> OTHER INFORMATION: n=a,c,g or t
44 <220> FEATURE:
45 <221> NAME/KEY: unsure
46 <222> LOCATION: (262)
47 <223> OTHER INFORMATION: n=a,c,g or t
49 <220> FEATURE:
50 <221> NAME/KEY: unsure
51 <222> LOCATION: (308)
52 <223> OTHER INFORMATION: n=a,c,g or t
54 <220> FEATURE:
55 <221> NAME/KEY: unsure
56 <222> LOCATION: (312)
57 <223> OTHER INFORMATION: n=a,c,g or t
59 <400> SEQUENCE: 1

ENTERED

RAW SEQUENCE LISTING

DATE: 05/15/2002

PATENT APPLICATION: US/10/018,902

TIME: 16:07:11

Input Set : A:\BB1373 USPCT Corrected Seq List.txt

Output Set: N:\CRF3\05152002\J018902.raw

W--> 60 ccaggcgcag gccttgcaa aggtgccan cgtcgccgcc ctcgagcagt tcaagatata 60
 61 cgccgaccgg tacatgaagg aaaggagtag catagctgtg ataggcctca gtgtacacac 120
 62 agcaccagtg gagatggcgt gtaaaaactt gctgttgtag aggaactgtg gccccgagct 180
 63 attcaagaac ttactagcc tgaaccatat tgaagaggcg tgctgttgct tgagtgcact 240
 W--> 64 gtgattngaa ttgganaatt tnatgtggtg ggcgctatcc atgggaaccg tggttatcag 300
 W--> 65 agaaagtnag tn 312

68 <210> SEQ ID NO: 2
 69 <211> LENGTH: 63
 70 <212> TYPE: PRT
 71 <213> ORGANISM: Zea mays
 73 <220> FEATURE:
 74 <221> NAME/KEY: UNSURE
 75 <222> LOCATION: (10)
 76 <223> OTHER INFORMATION: Xaa = ANY AMINO ACID
 78 <220> FEATURE:
 79 <221> NAME/KEY: UNSURE
 80 <222> LOCATION: (46)..(47)
 81 <223> OTHER INFORMATION: Xaa = ANY AMINO ACID
 83 <400> SEQUENCE: 2

W--> 84 Gln Ala Gln Ala Leu Ala Lys Ala Ala Xaa Val Ala Ala Leu Glu Gln
 85 1 5 10 15
 87 Phe Lys Ile Ser Ala Asp Arg Tyr Met Lys Glu Arg Ser Thr Ile Ala
 88 20 25 30
 W--> 90 Val Ile Gly Leu Ser Val His Thr Ala Pro Val Glu Met Xaa Xaa Lys
 91 35 40 45
 93 Leu Ala Val Ala Glu Glu Leu Trp Pro Arg Ala Ile Gln Glu Leu
 94 50 55 60

97 <210> SEQ ID NO: 3
 98 <211> LENGTH: 1924
 99 <212> TYPE: DNA
 100 <213> ORGANISM: Zea mays
 102 <400> SEQUENCE: 3

103 ccacgcgtcc gcatcaataa agaggagctt gggaagttgc caaggcctcc gatttcgcta 60
 104 atgcgacgat aatggcgacc acgacgtcag cgaccaccgc cgccgcagca gccgccacca 120
 105 ccgccaaagcc gcgggggtcg tcgtcggccc tctgccagag ggtggccggc ggccggcaggc 180
 106 ggcgctccgg ggtggtgcgg tgcgacgcgc ccggcggtga ggcccaggcg caggccgtgg 240
 107 caaaggctgc cagcgtcgcc gccctcgagc agttcaagat atccgccgac cggtagatga 300
 108 aggaaggag tagcatagct gtgataggcc tcagtgtaca cacagcacca gtggagatgc 360
 109 gtgaaaaact tgctgttgca gaggaactgt ggcccgtgc tattcaagaa ctactagcc 420
 110 tgaaccatat tgaagaggct gctgttctta gtacctgtaa tagaatggaa atttatgtgg 480
 111 tggcgctatc atgggaaccg ggtatcagag aagtagtgga ctggatgtcg aagaaaagtg 540
 112 gtattcccgcc ttccgagctt agggagcacc tgttcatctt gcgaagcagt gatgccacac 600
 113 gccatctgtt tgagggtgtca gctggccttg actctttggt tctcggtgaa ggacaaatcc 660
 114 ttgctcaggt taaacaagtt gtgaggagtg gacagaacag tggaggcttg ggaaagaaca 720
 115 tcgataggat gttcaaggat gcaatcactg ctggaaagcg tgtccgcagc gagaccaaca 780
 116 tatcatctgg tgctgtttct gtcagttcag cggcggttga actggccctg atgaagcttc 840
 117 cgaagtctga agcactgtca gctaggatgc ttctgattgg tgctggtaaa atgggaaagc 900
 118 tagtgatcaa acatctggtt gccaaaggat gcaagaaggt tgttgtggtg aaccgctccg 960
 119 tggaaagggt ggatgctatt cgtgaggaga tgaaagatat agagatcgtg tacaggcctc 1020

RAW SEQUENCE LISTING

DATE: 05/15/2002

PATENT APPLICATION: US/10/018,902

TIME: 16:07:11

Input Set : A:\BB1373 USPCT Corrected Seq List.txt

Output Set: N:\CRF3\05152002\J018902.raw

```

120 tctcagacat gtatcaagct gctgctgaag ctgatgtcgt gttcaccagc accgcatctg 1080
121 aaacttcatt gttcgcaaaa gaacacgcag aggcactccc ccctgtctct gatactatgg 1140
122 gaggtgttcg cctggtttgtc gacatatctg tccccaggaa tgtcagcgca tgtgtgtctg 1200
123 aagttggcgc tgcacgagtg tacaatgtcg acgacttgaa agaggtggtg gaagccaaca 1260
124 aggaggaccg gctcaggaaa gcaatggagg cgcagacaat catcaccgaa gaactgagac 1320
125 ggttcgaggc atggagggac tcgctggaga ccgttccgac catcaagaag ctgaggtcgt 1380
126 acgcgagacg gatcagggcc tcggagctcg agaagtgcct gcagaaagta ggtgaggacg 1440
127 ccctcaccaa gaagatgagg agagccatcg aggagctgag caccggcatc gttaacaagc 1500
128 tcctccatgg cccgtgcag cacctgaggt gcgacggcag cgacagccgc acccttgacg 1560
129 agacgctcga gaacatgcac gccctcaacc ggatgttcag cctcgacatg gagaaggcga 1620
130 tcatcgagca gaagatcaag gccaaaggtgg agaagacaca aaactgaggc caggaagcaa 1680
131 tttttctacc accattatct atatatatag cgtctccaat ctcatccat ttttttatcc 1740
132 tttcactcag tgagcccttc ccctgtcac tgtgatcgtt aactgtgtct gtgaattaga 1800
133 gccatggcag cgtgttgtca ataacagcaa tgtgtcccaa ttccccacag aagaaagact 1860
134 atatttatat gcatttattg gagcaaatag tttacttaaa aaaaaaaaaa aaaaaaaaaa 1920
135 aaag 1924
138 <210> SEQ ID NO: 4
139 <211> LENGTH: 531
140 <212> TYPE: PRT
141 <213> ORGANISM: Zea mays
143 <400> SEQUENCE: 4
144 Met Ala Thr Thr Thr Ser Ala Thr Thr Ala Ala Ala Ala Ala Thr
145 1 5 10 15
147 Thr Ala Lys Pro Arg Gly Ser Ser Ser Ala Leu Cys Gln Arg Val Ala
148 20 25 30
150 Gly Gly Gly Arg Arg Arg Ser Gly Val Val Arg Cys Asp Ala Ala Gly
151 35 40 45
153 Val Glu Ala Gln Ala Gln Ala Val Ala Lys Ala Ala Ser Val Ala Ala
154 50 55 60
156 Leu Glu Gln Phe Lys Ile Ser Ala Asp Arg Tyr Met Lys Glu Arg Ser
157 65 70 75 80
159 Thr Ile Ala Val Ile Gly Leu Ser Val His Thr Ala Pro Val Glu Met
160 85 90 95
162 Arg Glu Lys Leu Ala Val Ala Glu Glu Leu Trp Pro Arg Ala Ile Gln
163 100 105 110
165 Glu Leu Thr Ser Leu Asn His Ile Glu Glu Ala Ala Val Leu Ser Thr
166 115 120 125
168 Cys Asn Arg Met Glu Ile Tyr Val Val Ala Leu Ser Trp Asn Arg Gly
169 130 135 140
171 Ile Arg Glu Val Val Asp Trp Met Ser Lys Lys Ser Gly Ile Pro Ala
172 145 150 155 160
174 Ser Glu Leu Arg Glu His Leu Phe Ile Leu Arg Ser Ser Asp Ala Thr
175 165 170 175
177 Arg His Leu Phe Glu Val Ser Ala Gly Leu Asp Ser Leu Val Leu Gly
178 180 185 190
180 Glu Gly Gln Ile Leu Ala Gln Val Lys Gln Val Val Arg Ser Gly Gln
181 195 200 205
183 Asn Ser Gly Gly Leu Gly Lys Asn Ile Asp Arg Met Phe Lys Asp Ala
184 210 215 220

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/018,902

DATE: 05/15/2002

TIME: 16:07:11

Input Set : A:\BB1373 USPCT Corrected Seq List.txt

Output Set: N:\CRF3\05152002\J018902.raw

```

186 Ile Thr Ala Gly Lys Arg Val Arg Ser Glu Thr Asn Ile Ser Ser Gly
187 225 230 235 240
189 Ala Val Ser Val Ser Ser Ala Ala Val Glu Leu Ala Leu Met Lys Leu
190 245 250 255
192 Pro Lys Ser Glu Ala Leu Ser Ala Arg Met Leu Leu Ile Gly Ala Gly
193 260 265 270
195 Lys Met Gly Lys Leu Val Ile Lys His Leu Val Ala Lys Gly Cys Lys
196 275 280 285
198 Lys Val Val Val Val Asn Arg Ser Val Glu Arg Val Asp Ala Ile Arg
199 290 295 300
201 Glu Glu Met Lys Asp Ile Glu Ile Val Tyr Arg Pro Leu Ser Asp Met
202 305 310 315 320
204 Tyr Gln Ala Ala Ala Glu Ala Asp Val Val Phe Thr Ser Thr Ala Ser
205 325 330 335
207 Glu Thr Ser Leu Phe Ala Lys Glu His Ala Glu Ala Leu Pro Pro Val
208 340 345 350
210 Ser Asp Thr Met Gly Gly Val Arg Leu Phe Val Asp Ile Ser Val Pro
211 355 360 365
213 Arg Asn Val Ser Ala Cys Val Ser Glu Val Gly Ala Ala Arg Val Tyr
214 370 375 380
216 Asn Val Asp Asp Leu Lys Glu Val Val Glu Ala Asn Lys Glu Asp Arg
217 385 390 395 400
219 Leu Arg Lys Ala Met Glu Ala Gln Thr Ile Ile Thr Glu Glu Leu Arg
220 405 410 415
222 Arg Phe Glu Ala Trp Arg Asp Ser Leu Glu Thr Val Pro Thr Ile Lys
223 420 425 430
225 Lys Leu Arg Ser Tyr Ala Asp Arg Ile Arg Ala Ser Glu Leu Glu Lys
226 435 440 445
228 Cys Leu Gln Lys Val Gly Glu Asp Ala Leu Thr Lys Lys Met Arg Arg
229 450 455 460
231 Ala Ile Glu Glu Leu Ser Thr Gly Ile Val Asn Lys Leu Leu His Gly
232 465 470 475 480
234 Pro Leu Gln His Leu Arg Cys Asp Gly Ser Asp Ser Arg Thr Leu Asp
235 485 490 495
237 Glu Thr Leu Glu Asn Met His Ala Leu Asn Arg Met Phe Ser Leu Asp
238 500 505 510
240 Met Glu Lys Ala Ile Ile Glu Gln Lys Ile Lys Ala Lys Val Glu Lys
241 515 520 525
243 Thr Gln Asn
244 530
247 <210> SEQ ID NO: 5
248 <211> LENGTH: 510
249 <212> TYPE: DNA
250 <213> ORGANISM: Oryza sativa
252 <220> FEATURE:
253 <221> NAME/KEY: unsure
254 <222> LOCATION: (326)
255 <223> OTHER INFORMATION: n=a,c,g or t
257 <220> FEATURE:

```

RAW SEQUENCE LISTING

DATE: 05/15/2002

PATENT APPLICATION: US/10/018,902

TIME: 16:07:11

Input Set : A:\BB1373 USPCT Corrected Seq List.txt

Output Set: N:\CRF3\05152002\J018902.raw

258 <221> NAME/KEY: unsure
259 <222> LOCATION: (335)
260 <223> OTHER INFORMATION: n=a,c,g or t
262 <220> FEATURE:
263 <221> NAME/KEY: unsure
264 <222> LOCATION: (344)
265 <223> OTHER INFORMATION: n=a,c,g or t
267 <220> FEATURE:
268 <221> NAME/KEY: unsure
269 <222> LOCATION: (355)
270 <223> OTHER INFORMATION: n=a,c,g or t
272 <220> FEATURE:
273 <221> NAME/KEY: unsure
274 <222> LOCATION: (362)
275 <223> OTHER INFORMATION: n=a,c,g or t
277 <220> FEATURE:
278 <221> NAME/KEY: unsure
279 <222> LOCATION: (364)
280 <223> OTHER INFORMATION: n=a,c,g or t
282 <220> FEATURE:
283 <221> NAME/KEY: unsure
284 <222> LOCATION: (371)
285 <223> OTHER INFORMATION: n=a,c,g or t
287 <220> FEATURE:
288 <221> NAME/KEY: unsure
289 <222> LOCATION: (378)
290 <223> OTHER INFORMATION: n=a,c,g or t
292 <220> FEATURE:
293 <221> NAME/KEY: unsure
294 <222> LOCATION: (382)
295 <223> OTHER INFORMATION: n=a,c,g or t
297 <220> FEATURE:
298 <221> NAME/KEY: unsure
299 <222> LOCATION: (390)
300 <223> OTHER INFORMATION: n=a,c,g or t
302 <220> FEATURE:
303 <221> NAME/KEY: unsure
304 <222> LOCATION: (399)
305 <223> OTHER INFORMATION: n=a,c,g or t
307 <220> FEATURE:
308 <221> NAME/KEY: unsure
309 <222> LOCATION: (403)
310 <223> OTHER INFORMATION: n=a,c,g or t
312 <220> FEATURE:
313 <221> NAME/KEY: unsure
314 <222> LOCATION: (411)..(412)
315 <223> OTHER INFORMATION: n=a,c,g or t
317 <220> FEATURE:
318 <221> NAME/KEY: unsure

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 05/15/2002
PATENT APPLICATION: US/10/018,902 TIME: 16:07:12

Input Set : A:\BB1373 USPCT Corrected Seq List.txt
Output Set: N:\CRF3\05152002\J018902.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 30,247,256,262,308,312
Seq#:2; Xaa Pos. 10,46,47
Seq#:5; N Pos. 326,335,344,355,362,364,371,378,382,390,399,403,411,412,434
Seq#:5; N Pos. 444,448,453,483,490,492,494,502
Seq#:6; Xaa Pos. 76
Seq#:9; N Pos. 217,241,243,301,360,373,405,412,426,439,447,515
Seq#:13; N Pos. 496,500
Seq#:17; N Pos. 1,2,5,8,16,18,21,22,27,33,35,40,101,232,298,313,349,360,377
Seq#:17; N Pos. 378,384,388,391,392,397,400,407,410,423,424,425,426,427,428
Seq#:17; N Pos. 431,432,434,440,441,446,451,454,455
Seq#:18; Xaa Pos. 8,21
Seq#:23; N Pos. 136,220,266,334,341,348,353,356,360,382,385,396,404,410,416
Seq#:23; N Pos. 419,434,451,453
Seq#:24; Xaa Pos. 32,60,75,80,98
Seq#:26; Xaa Pos. 322
Seq#:27; N Pos. 321,334,350,356,362,367,375,400,402,417,439,460,464,467,475
Seq#:27; N Pos. 490,499,507,528,530,537,602,603,609,625,636,650